DOCUMENT RESUME

ED 077 972 TM 002 781

AUTHOR Popham, W. James
TITLE Applications of Teaching Performance Tests to Inservice and Preservice Teacher Education.
PUB DATE 73
NOTE 22p.; Paper presented at annual meeting of American Educational Research Association (New Orleans, Louisiana, February 25-March 1, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Annotated Bibliographies; *Evaluation Techniques; Higher Education; *Performance Tests; *Program Evaluation; Speeches; *Teacher Education; *Teaching Skills; Test Construction

ABSTRACT

Teacher educators have been plagued with the problem of devising useful instructional interventions, but often have been unable to assess the quality of their effcrts because of the unavailability of satisfactory criterion measures. The application of teaching performance tests as (1) an instructional intervention, and as (2) a formative or summative evaluation criterion in connection with inservice and preservice teacher education programs is described. Applications of the performance test strategy to both types of programs are presented. This measurement strategy has utility for preservice credential programs and inservice staff development enterprises. (Author)

ED 077972

TM 002 281

A paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 26-March 1, 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

# APPLICATIONS OF TEACHING PERFORMANCE TESTS TO INSERVICE AND PRESERVICE TEACHER EDUCATION

W. James Popham
University of California, Los Angeles
and
The Instructional Objectives Exchange

Irrespective of whether the programs are referred to as "performance-based," "competency-based," or "skill-focused," the advocacy of teacher preparation schemes designed to promote measurable capabilities of instructors is beginning to be quite fashionable among teacher educators. A recent survey[1] of California teacher education institutions found that two-thirds of the institutions participating in the survey indicated they were engaged at least partially in competency-based teacher preparation. There is even a multistate consortium[2] on performance-based teacher education featuring such forward-looking states as Arizona, Florida, Minnesota, New York, Oregon, Texas, Utah, Vermont and Washington. Surely, if most teacher educators are not caught up in the competency-based game, then we are at least witnessing the pre-game warmups.

Whether competency-based teacher education will become something other than one of those ephemeral fads so common in teacher education remains to be seen. Although many teacher educators are quick to join the competency-based movement at the verbal level, few are willing to devote the requisite energy to devising the criterion measures without which the approach is only rhetoric. In the recent survey[3] of California teacher education colleges and universities, no institution reported satisfactory assessment procedures for a competency-based program. As with so many educational innovations, advocacy is less expensive than workable implementation procedures.

## Teacher Competencies as Enabling Skills

But there is another aspect of the competency-based teacher education movement which is equally intriguing. If you sit in on almost any discussion among proponents of competency-based teacher

---

[1] Performance-Based Teacher Education, Vol. 1, No. 5, December 1972, pp. 1, 5.

[2] Multistate consortium on performance-based teacher education, c/o Bureau of Teacher Education, New York State Education Department, Albany, N.Y.

[3] op. cit.

FILMED FROM BEST AVAILABLE COPY

education, you will hear them describing the teacher competencies they are trying to promote as though they were ends in themselves. As with many recent religious converts, the fervor of these teacher educators for competency identification has become so all-consuming that they lose sight of what the competencies are really supposed to accomplish.

Reducing the problem to its essentials, we can see that whatever competencies a teacher acquires must be viewed as vehicles for making that teacher more effective. And more effective, in this instance, means a teacher better able to help learners. Thus, the competencies most frequently identified by performance-prone teacher educators are really en route skills which should contribute to the terminal skill of being able to help learners. For example, suppose one of the teacher competencies we are trying to promote involves the teacher's ability to view real or simulated instructional situations and identify the extent to which certain instructional tactics have been employed. We assume that the teacher who can master such a skill will subsequently be able to apply this skill in real instructional situations. Such skills should thus be viewed as precursive to one's becoming an effective teacher.

Now the point of this distinction between en route and terminal competencies is that the bulk of competencies currently viewed as the staples of performance-based teacher education are well removed from those which might legitimately be viewed as terminal. And unless en route skills are constantly verified as being actual contributors to terminal skills, then we have little assurance that defensible competencies are being promoted.

Some of the competencies sought by teacher educators should be closer to the terminal proficiencies we wish teachers to display. The purpose of this paper is to explore the teacher education applications of one measurement approach designed to assess such a near-terminal competency. The measurement approach under consideration is the teaching performance test.

## Teaching Performance Tests: Description and Rationale

In brief[4], teaching performance tests work as follows: An instructor is presented with one or more explicit instructional objectives (plus a sample test item) and is directed to prepare a short lesson designed to accomplish the objective(s). If the objective deals with a topic presumed to be unfamiliar to the teacher, then relevant background information is made available. After planning the lesson, the teacher instructs a group of learners (either children or adults) for a short period of time, e.g., 15 minutes. The number of learners can be as few as a half dozen or as many as an entire class. At the conclusion of

---

[4]McNeil, J. D., and Popham, W.J., "The Assessment of Teacher Competence," Chapter 7, The Handbook of Research on Teaching, R.M.W. Travers (Ed.), MacMillan, in press.

the lesson the learners are given a posttest[5] based on the objective. While not previously seen by the teacher, the nature of the test is readily inferrable from the objective the teacher has been attempting to achieve. Characteristically, learners are also asked to rate how interesting the lesson was. If such an interest rating is employed, the teacher is apprised of the forthcoming rating and encouraged to design a lesson which not only accomplishes the cognitive objective, but also promotes positive learner interest ratings.

Using these two indicators, an estimate is provided of the teacher's ability to promote prespecified objectives, both cognitive (as reflected by posttest performance) and affective (as reflected by the interest ratings). Now it may be argued that an unrepresentative estimate of instructional prowess is yielded by measuring a teacher's ability to promote learner attainment of prespecified objectives for a small group of learners during a brief time period. Yet, while perhaps not as representative of the real teaching world as we might wish, there are sufficient parallels with reality that such an assessment procedure may have utility for teacher educators. In particular, since it more closely approximates a terminal teaching skill than many of the competencies currently being fostered by performance-based teacher educators, it may have advantageous instructional and evaluational dividends. The remainder of this analysis will (1) set forth three distinctive applications of teaching performance tests in teacher education operations, (2) describe actual inservice and preservice situations in which performance tests have been employed, (3) identify usage guidelines derived from these experiences, and (4) discuss certain problems which have arisen in the use of teaching performance tests.

## Application One: A Focusing Mechanism

It is the writer's belief that much of the educational ineffectiveness which exists in our schools can be attributed directly to teachers' preoccupations with instructional process. Far too many teachers are caught up with concerns about devising new and exciting ways of teaching, without ever verifying what effects those procedures have on children. Some teachers pridefully announce that they strive to "teach their class differently every year," never recognizing that they may be abandoning one year an approach that was truly effective the previous year. Innovations are adulated for their own sake. For instance, open schools are currently in vogue. Ten years ago it was nongraded schools. A decade earlier we were praising the raptures of the core curriculum.

Not that there is anything intrinsically wrong with these new instructional approaches, for surely they possess many meritorious features. It's just that too many educators succumb to

---

[5]If novel subject matter is employed, no pretest is typically employed. With less esoteric topics, a pretest may be utilized to identify sufficiently naive learners.

the lure of an attractive instructional process without checking the quality of its impact on learners. And that, after all, should be the reason we search for better instructional procedures.

Hence, as a method of counteracting what appears to be an almost hereditary concern about instructional process, frequent use of teaching performance tests can provide a mechanism to focus the teacher's attention on the effects of instruction. Since in a performance test situation the quality of a teacher's efforts is predicated on results achieved with learners, both cognitive and affective, it is difficult to discount the effects of instruction on pupils. For example, if you are a prospective teacher who during a semester is obliged to teach a half dozen or more mini-lessons (as short duration teaching performance tests are sometimes called), and the first concern after your lesson is a determination of its effects on learners, it is difficult to see how you would not soon begin to view as important a lesson's impact on pupils.

This initial application, therefore, is instructional in nature. More specifically, it is designed to foster a disposition on the part of the teacher, namely, a disposition to view as important the effects of instruction on learners.

Application Two: A Setting for Testing the Value of Instructional Tactics

Even though the foregoing application, i.e., as a mechanism for focusing teachers' attention on the consequences of instruction, may have suggested that attention to instructional procedures was somehow reprehensible, such is surely not the case. We can only secure good results with learners if we use appropriate instructional processes. The trick is to apply instructional techniques judiciously in such a way that we can either verify their efficacy (in terms of effects on learners), or at least be able to make high probability guesses that a given technique will yield desirable results with pupils.

A second application of teaching performance tests involves their use as a method of allowing teachers to test the differential effectiveness of various instructional techniques. Teachers can complete a series of minilessons attempting to incorporate different instructional tactics, then judge their worth in terms of the results yielded with learners. For instance, suppose a teacher taught the same minilessons to two different groups of comparable learners, the lesson being essentially the same except that one lesson provided much opportunity for learners to practice the skills called for in the minilesson's objective, while the other lesson provided no such practice. The teacher could then contrast the posttest results of both lessons and begin to reach a conclusion regarding, for certain kinds of instructional objectives, the efficacy of providing relevant practice.

Not that a tactic-present versus tactic-absent design must be employed in this second application of performance tests, for a teacher can often gain insights regarding the value of a given

instructional procedure from using the procedure even without the "control" treatment. This is particularly true when for a particular performance test there are some normative data which, even in rough terms, yield an estimate of how well teachers typically perform on the lesson. This point will be treated in more detail later.

An important aspect of this application of teaching performance tests is that a teacher need not be the actual instructor in a minilesson to profit from the minilesson teacher's experience. A typical format for the conduct of minilessons, either those taught to younger learners or to a group of one's peers, is the post-lesson analysis session. During this session, the teacher's instructional approach is appraised in terms of its effects on learners. Many teachers report they learn as much from watching the minilesson teacher's lesson, then analyzing its strengths or shortcomings, as they do from actually teaching the lesson themselves.

Frequent teacher performance tests, either for teachers as minilesson instructors or as an object for group analysis of other's instructional efforts, can provide the focus of a consequence-oriented preservice or inservice teachers education program.

Application Three: A Formative or Summative Program Evaluation Device

Developmental work with teaching performance tests is still at such an early stage that it may be imprudent to employ them for the evaluation of individual teachers. The only exception might be for isolating instructors who are extremely weak or strong in their ability to accomplish prespecified goals. Nonetheless, as a program evaluation assessment technique, performance tests may have considerable utility. Indeed, the most important use of teaching performance tests may be as instruments to aid in the appraisal of inservice or preservice teacher education enterprises.

The argument, briefly, is that if a teacher education program sets out to promote teachers' abilities to accomplish prespecified objectives, then the program can be legitimately evaluated in terms of its ability to do so. Here's how the evaluation strategy might work. At the outset of a teacher education program, e.g., a preservice credential sequence or an extended staff development institute, a representative sample (or all) of the participating teachers could complete one of two different performance tests (e.g., Test X and Test Z, with half the teachers completing Test X and half Test Z). At the close of the program, the teachers would complete the other performance test. The prediction would be that for both tests the teachers' post-program efforts would produce markedly better results than the pre-program efforts. As a summative evaluation stratagem, such an approach could yield devastating information. What happens, for instance, if a teacher education program discovers its teachers are no better able after intensive instruction to promote learner mastery of objectives than they were prior

to instruction? Surely drastic changes in the program seem warranted. Perhaps, if such failures are recurring phenomena, the program should be terminated.

If performance tests are employed at earlier stages of the program, with a view to guiding program modifications during the course of the instruction, then formative evaluation benefits can also be derived from the use of teaching performance tests.

In review, then, we have briefly examined three possible applications of teaching performance tests in connection with either preservice or inservice teacher education efforts. There will, of course, be other uses of performance tests in teacher education. For example, a professor might use teaching performance tests as a motivating mechanism, showing novice students that more skilled teachers can out-perform beginners in such instructional situations. But focusing for the moment on the three application strategies described thus far, we can examine some actual utilizations of teaching performance tests.

## Actual Inservice and Preservice Applications

The Jordan Complex. One of the first reported uses of teaching performance tests to aid practicing teachers took place during 1971 in the Jordan Complex, an affiliated group of urban schools in the Los Angeles City School District. Under the leadership of LaVerne Parks, Complex Director, groups of teachers met weekly after regular school hours to witness each other taking turns teaching minilessons, then discussing the merits of the teaching approaches. Lessons were divided about evenly between those designed for young children (volunteer pupils were used as learners) and those designed for adults (teachers participating in the program took turns serving as learners).

Reactions to the program, extending over several months, were quite positive. Merle Williamson and Joyce Cooper, the class leaders, gathered anonymously-supplied course evaluation data indicating that 84 per cent of the participating teachers felt the class had helped them (1) understand the role of instructional objectives, (2) plan lessons for given periods of time, (3) develop alternative instructional strategies, and (4) critique lessons on the basis of post-test results. Eighty-seven per cent of the participants reported that the class had helped them personalize instruction. Mrs. Parks remarked, in reviewing the experience, that "One of the most exciting outcomes from this course was that teachers began to look more critically at themselves and their peers in terms of factors contributing to successful pupil results."[6]

UCLA. As might be anticipated because of the writer's affiliation, use of teaching performance tests in the teacher education program at UCLA over the past years has been fairly extensive. All three of the applications discussed in the previous section have

---

[6]Personal communication to the writer, May 15, 1972.

been employed, with attention generally given to the use of performance tests as an instructional intervention.

Typically, teaching performance tests have been used with pre-service candidates, usually involving lessons taught to other members of the class. The setting for these lessons is ordinarily referred to as a minilesson clinic and features the customary teaching-testing-analysis model. Usually one or two minilessons are taught during a two-hour clinic session. During some terms we have tried to squeeze three minilesson assignments into a single two-hour lab period. On other occassions we have required the prospective teachers to generate their own objectives and tests, then try to accomplish the objectives in a short-term lesson, as an additional exercise in promoting learner goal attainment.

The most recent procedure we have employed for using teaching performance tests is described in some detail in the Appendix. Briefly, it involves the use of weekly nine-student minilesson clinics during which one preservice credential candidate teaches six classmates while two classmates plus a teaching assistant serve as instructional analysts. In addition, each credential candidate is obliged to teach at least three minilessons outside of class time to small groups of adults. Thus, in a ten week academic quarter, prospective teachers have about two dozen opportunities to serve as minilesson teachers, analysts, or learners.

As our use of minilesson clinics has increased, there has naturally been great interest in the manner in which the teacher education students were receiving them. At the end of the Fall quarter, a quarter during which minilessons were employed according to the scheme presented in the Appendix, students were asked to supply anonymous evaluations of the course at its conclusion. Little structure in the evaluation form was presented to the students, only requests to isolate parts of the course they liked most, liked least, etc. Of the 58 students who mentioned the minilesson clinics, 32 were positive and 26 were negative. In view of the fact that none of the minilesson clinic leaders possessed any prior experience with minilesson clinics, either as participants or supervisors, hence were probably less skillful as clinic leaders than might have been wished, these results are not distressing.

One of the hopes in setting up minilesson assignments so that students were obliged to teach the same minilesson a second time (outside of class time) was that the minilesson teacher would profit from the clinic critique session. Hopefully, the insights gained from the analysis of the teacher's lesson would lead to instructional improvements when the minilesson was retaught and a comparable posttest form was used to assess learner achievement.

Happily, this appears to have been the case in the Fall, 1972 UCLA situation for, from the reports supplied by class members (to which no grade credit was applied), both cognitive and

affective performances increased more frequently on the retaught lessons than they decreased, as can be seen in Table 1. Of course, these data are only suggestive in view of the fact that there were no controls exercised over the ability of learners in the retaught or initially taught lessons.

Table 1. Results of minilessons retaught after initial minilesson clinic analysis

| | First lesson Superior | No Difference | Second Lesson Superior |
|---|---|---|---|
| Cognitive Posttest | n = 18 | n = 6 | n = 31 |
| Interest Rating | n = 13 | n = 9 | n = 33 |

Recently we have also used performance tests in connection with program evaluation. During the Fall, 1972 quarter, six minilessons were randomly assigned to approximately a dozen credential candidates (from a particular teacher education course) at the beginning of the quarter, and to a different group of credential candidates at or near the close of the quarter. In gross terms, the prediction was that the performance of these prospective teachers would be better later in the quarter, presumably after the impact of the teacher education program had worked its "beneficial" effects.

Since the minilessons were to be taught to classmates, care had to be taken not to involve any students in minilessons at the first of the quarter (either as teachers or learners) which they would encounter at the close of the quarter. Since six different minilessons were used, this posed no problem. Although it had been planned to have three teachers per minilesson on both an early-in-course and late-in-course basis, with student attrition, missed assignments, etc., a less balanced performance distribution resulted. The minilessons were taught to groups of approximately six learners (according to the procedural scheme described in the Appendix), with the early-in-course minilessons occuring during the second week on the quarter. Ideally, the first week would have been used for the pretest minilesson, but course organizational requirements dictated that the second week was a more reasonable choice. Minilessons completed during weeks eight, nine, or ten of the ten-week quarter constituted the late-in-course measures. Since all minilessons were completed on Fridays, this means that prior to the late course minilessons the credential candidates had typically experienced a minimum of about eight weeks of instruction.

The teacher education program variation under consideration involved daily two-hour class sessions, including lectures, discussions, observations in public schools, and the minilesson clinic activities.

Results of the early course versus late course measurement are presented in Table 2 for both the cognitive measure (posttest percentage correct) and affective measure (interest rating). All data included in Table 2 are mean results for a given minilesson teacher, typically based on an n of five or six classmate-learners. It should be noted that since six different minilessons were employed, with different levels of difficulty and interest associated with each, interpretation of the data should focus on the columns of the table.

Table 2. Mean results of early-in-course and late-in-course minilessons taught by credential candidates in a UCLA teacher education program.

| | Minilessons | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Cognitive Results | Posttest Percentages Correct | | | | | |
| Early-in-Course | 72<br>74<br>80 | 90 | 36 | 84<br>90<br>96 | 82 | 60<br>74 |
| Late-in-Course | 88<br>90<br>92 | 73<br>95 | 36<br>50<br>65<br>67 | 86<br>97 | 82 | 82 |
| | Minilessons | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Affective Results | Ratings of Lesson's Interest (5 = hi, 1 = low) | | | | | |
| Early-in-Course | 2.7<br>3.2<br>3.4 | 3.8 | 3.8 | 3.6<br>3.8<br>4.0 | 3.0 | 2.6<br>3.8 |
| Late-in-Course | 3.0<br>3.8<br>4.0 | 3.3<br>4.5 | 3.2<br>5.0<br>4.1 | 4.0 | 3.2 | 4.6 |

An examination of Table 2 will reveal that the predicted results were supported by the data. For both indices, i.e., mean posttest percentages correct and interest ratings, the late-in-course performances exceeded the early-in-course performances.

By converting the results for each performance test to standard scores for that measure, with a mean of 50 and a standard deviation of 10, it was then possible to pool the data from all six performance tests and compute separate t tests for both the cognitive and affective results. A summary of these analyses is presented in Table 3 where it can be seen that a significant difference was present in both instances, with a mean difference of .83 standard deviation unit present in the case of the cognitive measure and .97 standard deviation unit for the affective measure. Both differences, as hoped, favored the late-in-course teaching performance tests.

Table 3. Analysis of early-in-course and late-in-course minilesson results, pooled on the basis of standard score transformations.

| | n | $\bar{X}$ | S.D. | $\bar{X}_{diff.}$ | t | p* |
|---|---|---|---|---|---|---|
| Cognitive Results | | | | | | |
| Early-in-course | 11 | 45.46 | 7.86 | 8.31 | 2.25 | <.025 |
| Late-in-course | 13 | 53.77 | 9.28 | | | |
| Affective Results | | | | | | |
| Early-in-course | 11 | 44.73 | 5.59 | 9.73 | 2.57 | <.01 |
| Late-in-course | 13 | 54.46 | 10.90 | | | |

*one-tailed

This represents the initial time, at least to the writer's knowledge, that teaching performance tests have been used in this manner as a teacher education program evaluation technique. It is obvious, even from examining the data table, that some refinements are in order, e.g., the disparate n's in the columns, etc. Nevertheless, as an illustration of the use of performance tests for program evaluation purposes, the foregoing description may be of some utility.

California State University, Northridge. During summer, 1972 two inservice workshops for approximately 50 teachers and administrators were sponsored by California State University, Northridge in coordination with Administrative Area K of the Los Angeles City Schools. Clare Rose, the instructor for both workshops, reports that teaching performance tests were used in the workshops as a vehicle for achieving one of the workshop goals, namely, that participants would be able to supervise objectives-based instructional improvement programs.

A demonstration minilesson was presented in each workshop followed, one week later, by having all workshop participants serve either as minilesson teachers or students as particular minilessons were taught twice by six volunteers. Prior to the second teaching of each minilesson, an analysis of the initial lesson was carried out. Professor Rose reports that 90 per cent of the minilesson teachers were able to promote higher posttest results on the second lesson. She indicated that 100 per cent of the workshop participants reported, on anonymous end-of-session evaluation forms, that the minilesson activities had been valuable.

Professor Martin Levine has also employed teaching performance tests in preservice teacher education for several years at California State University, Northridge. Most recently, Professor Levine reported the following format for his use of performance tests:

> My preservice secondary education methods course meets on the campus of a participating junior high school, usually for three hours weekly (e.g., Wednesday 9-12). One or two classes of secondary pupils are assigned to the college methods class by the principal. Usually pupils in these classes are classified as "low ability learners" in need of more individual attention. Pupils are assigned at random to college trainees who are responsible for achieving prespecified instructional objectives issued by the instructor. Each trainee teaches his miniclass of three pupils for one hour during each weekly meeting of the course. Objectives deal with general study skills and thus are appropriate for both pupils who need to master this kind of objective and for college trainees who come from all of the different academic areas commonly taught in secondary schools. Trainees have a week to prepare to teach each objective. A teaching time limit of thirty minutes or less is usually set. The instructor administers a pretest and posttest. Trainees analyze their teaching effectiveness in terms of how well their pupils achieve the objectives. Trainees may re-teach objectives during subsequent meetings in cases which warrant it. From time to time, trainees combine miniclasses, with one member teaching to a prespecified objective while his peers observe the lesson for use of such instructional principles as practice and feedback. A postobservation conference is held immediately after the lesson. Usually instruction takes place in a large area such as the oral arts room or the school cafeteria where the instructor can monitor the entire process assisted by the master teachers.

## User Guidelines

Based on our limited experience to date, there appear to be a few guidelines which might be of value to teacher educators

considering the utilization of teaching performance tests. Some general suggestions regarding the use of performance tests are available elsewhere.[8]

First, it has become apparent that the capability of the minilesson analyst (when group use of performance tests is involved) is far more critical than had been anticipated. Unless the person supervising the analysis session is both convinced of the value of the activity and able to provide instructional insights (when, for example, poor learner performance occurs), then the minilesson may be far less profitable than possible. Too many teachers will write off minilesson teaching efforts, particularly if poor performances occur, as unrepresentative of their performance in a "real" teaching situation. Thus, one of the supervisor's missions is to clarify the parallels of the minilesson activity and regular classroom teaching. Further, a deft instructional analysis can provide an unsuccessful minilesson teacher with a promise of future success during re-taught or other subsequent lessons. The minilesson supervisor must eliminate the frustration that follows failure if no improvement plan is presented. Skilled supervisors are so important that, in the writer's view, unless they are available in sufficient numbers (or can be trained in time) group-type minilesson activities should not be undertaken.

A second point relates to the use of minilessons with children or adults. Reports to date suggest that while a steady diet of minilessons for young learners is palatable, exclusive use of adult learners (e.g., teacher educator classmates or colleagues) is less acceptable. Teacher educators who, for a variety of practical reasons, may prefer to rely on minilessons for adult learners, should strive to provide, as a change of pace, a few minilessons involving younger learners.

Third, there seem to be some discernable dividends associated with providing some type of normative data, even roughly displayed, against which to interpret one's performance as a minilesson teacher. Without such comparative data, the teacher or supervisor is hard put to tell whether a given performance is good or bad. Referring back to Table 2, it can be seen that there are clear differences, both in difficulty and probable interest, in certain of the minilessons. How is the minilesson teacher to know whether his/hers was a well designed and excecuted lesson if no interpretive framework is provided? We need to supply minilesson teachers with what any golfer needs to make the game more meaningful, a rough notice of what constitutes a par performance.[9]

---

[8]In the teaching improvement kits distributed by Instructional Appraisal Services, Box 24821, Los Angeles, California 90024.

[9]In the absence of clearly posted signs indicating pars (per hole) of three, four, or five, the writer's early experiences on a golf course would have suggested that an acceptable number of strokes per hole was something closer to 10, 15, or 20.

Of course, normative data are not absolutely indespensible, for when a minilesson teacher discovers that his or her learners are all scoring around 50 per cent on a task for which 90-100 per cent proficiency had been anticipated, then the teacher at least has the knowledge that expectations have not been attained -- and that's better than nothing. But comparative data are really a big help.

Problems

Teacher performance tests, when employed as resources for teacher education evaluation or instruction, are in their infancy. Not surprisingly, therefore, a score of diaper-related difficulties have already arisen. And until a definitive Dr. Spock volume arrives to deal with our dilemmas, we'll have to do some trying and, unfortunately, some erring.

One of the most basic problems facing would-be users of performance tests stems from the instruments themselves. We have not yet made an acceptable effort to delineate the defining dimensions of performance tests, in terms of their content, objectives, post-test nature, background information, difficulty level, etc. Almost all of the recently developed performance tests have been devised more or less on the basis of experience and intuition. This situation needs to be rectified without delay.

A related problem is the reliability of the teacher performance tests themselves. We do not yet have sufficient data to know how many minilessons a teacher must attempt before we can assess the teacher's overall level of competence on such tasks. It is expected that a teacher will not rate consistently high, say, on all minilesson attempts because his teaching score will depend in part on the subject matter of the lesson, in part on his own teaching approach to that lesson, and in part on how well other factors have been controlled. We are optimistic that teachers on the ends of the competency contimuum can be isolated for special assistance or special commendation.

A minor problem but one which can yield troublesome implementation difficulties, is the necessity to select topics for minilessons which will be viewed as important, both by the teachers who carry out the minilesson and by the students who are taught. In an effort to identify novel topics, thereby eliminating the need for pretesting, a few topics have been chosen which are so esoteric as to yield atypical (or all too typical) student apathy. Perhaps it may be wiser to select some main line curricular objectives and go to the trouble of pretesting prospective students in order to locate a suitably unknowledgeable learner group. Such minilesson topics might then be viewed as more meaningful by both teachers and learners.

Finally, the logistics problems associated with proper use of performance tests should be anticipated. An examination of the

step-by-step details given in the Appendix will reveal the level of organizational planning needed to head off confusion. For instance, it is highly desirable to provide teachers of unsuccessful minilessons with an opportunity to re-plan and then re-teach the lesson to a different group of learners. But providing these re-teaching opportunities takes a good deal of planning time. Faced with such planning frustrations, many teacher educators will be tempted to return to the less taxing, but perhaps less effective, lecture-discussion classroom format. Anticipating logistical problems can help avoid them. A competent secretary or teaching assistant can alleviate logistical distractions by working out the organizational requirements in advance.

Review

In retrospect, an attempt has been made in this paper to discuss possible applications of teaching performance tests to the activities of teacher educators. If a bias in favor of such applications was reflected in the paper, this was only natural, for such is the writer's bias. Although still a rather primitive tool, the performance test may be a valuable instrument to teacher educators. As we look at ancient man's hand axes, we may view them as incredibly simple devices, yet their impact was enormous. No strict analogy is being proposed here, only a plea to consider teaching performance tests as an additional tool in our teacher education kits. Think of how many sabretooth tigers we might slay.

Annotated Bibliography of
Teaching Performance Test Articles


Baker, E. L., "Relationship Between Learner Achievement and Instructional Principles Stressed During Teacher Preparation," Journal of Educational Research, November, 1969, 63(3):99-102. (A study is reported in which teaching performance tests are employed as a dependent measure in relationship to several independent instructional variables.)

_____., Teaching Performance Tests as Dependent Measures in Instructional Research, a paper presented at the annual meeting of the American Educational Research Association, New Orleans, February 25-March 1, 1973. (Techniques for using teacher performance tests in instructional research are described along with illustrative data typifying such applications.)

Belgard, M; Rosenshine, B. and Gage, N.L., "Explorations of the teacher's effectiveness in explaining, in Bellack, A.A. and Westbury, I." (Editors), Research into Classroom Processes: Recent Developments and Next Steps. New York, N.Y.: Teachers College Press, 1971. pp. 175-217. (Forty-three teachers taught two different 115 minute lessons to the same group of pupils. The correlation between residual class means was .47.)

Connor, A., Cross-validating Two Performance Tests of Instructional Proficiency. University of California, Los Angeles. USOE Project No. 8-I-174, 1969. (This investigation, conducted as part of a postdoctoral fellowship project by Connor, examines the correlation between the performance of 17 experienced teachers on two different performance tests. A low, positive relationship was reported.)

Glass, G. V., Statistical and Measurement Problems in Implementing the Stull Act, a presentation at the Stanford Invitational Conference on the Stull Act, October, 1972. (In this paper, Glass critiques a number of previous studies regarding teaching performance tests and offers the view that their current reliability is not adequate to warrant use in the evaluation of teachers.)

Justiz, T. B., A Method of Identifying the Effective Teacher. Doctoral thesis: University of California, Los Angeles, 1969. (Univ. microfilms No. 29-3022-A.) (This investigation assessed the degree of correlation between the performances of nine teachers on two different performance tests involving different pupils. A rank order correlation of .64 was reported. A second study was also described in which five teachers taught two different performance tests to the same students. A high rank order correlation was reported for the latter analysis.)

Levine, M. G., "The Effect on Pupil Achievement of a Criterion-referenced Instructional Model Used by Student Teachers," The Journal of Teacher Education., Vol. 23, No. 4, Winter 1972, pp. 477-481. (A 75 minute teaching performance test was used as a criterion measure to contrast the efficacy of alternative teacher education approaches.)

McNeil, J. D., "Performance Tests: Assessing Teachers of Reading," The Reading Teacher, 1972, pp. 622-627. (In this article a "teaching fair" is described in which prospective elementary teachers taught different groups of three elementary school pupils in consecutive teaching situations. The investigator reported a tendency for teachers to display relatively stable proficiency on such tasks.)

McNeil, J. D., and Popham, W. J., "The Assessment of Teacher Competence," Chapter 7, The Handbook of Research on Teaching, R.M.W. Travers (Ed.), MacMillan, in press. (This is a comprehensive review of research and evaluation strategies involving teacher effectiveness. Teaching performance tests are examined along with commonly employed teacher competency indicators according to various criteria for use in research and evaluation.)

Millman, J., Psychometric Characteristics of Performance Tests of Teaching Effectiveness. Paper presented to the American Educational Research Association, New Orleans, February, 1973. (Describes a number of small scale investigations carried out in Pasadena, California in 1971-72 in an effort to study the psychometric properties of teaching performance tests.)

_____., "Teacher Effectiveness: New Indicators for an Old Problem," Educational Horizons, Spring 1973. (A readable rationale statement endorsing the use of performance-based measures of instructor skill.)

Popham, W. J., Evaluating Instruction. Englewood Cliffs, New Jersey: Prentice-Hall. Inc., 1973. (Two chapters of this book are self-instruction programs designed to provide the reader with skills needed to construct and to use teaching performance tests.)

_____., "Performance Tests of Teaching Proficiency: Rationale, Development, and Validation," American Educational Research Journal. 1971, 8, 105-117. (This article describes the background of teaching performance tests and the development of such tests in the fields of social studies and vocational education during the mid-sixties. Attempts to validate the tests are reported and the validation strategy analyzed.)

_____., "Teaching Performance Tests as a Vehicle for Instructor Self-Evaluation," The National Elementary Principal, Vol. 52, No. 5, February, 1973. (This article describes alternative techniques of employing teaching performance tests as instruments for instructor self-appraisal.)

_____., "Teaching Skill Under Scrutiny," Phi Delta Kappan. Vol. 52, No. 10, June, 1971, pp. 599-602. (This article examines alternative techniques for assessing teacher skill and describes the possible role of teacher performance tests as an alternative to ratings, observations, and pupil performance on standardized tests.)

Appendix

TEAM S MINILESSON CLINIC GUIDELINES

UCLA Teacher Education Laboratory
Fall, 1972

Commencing with the second week of the quarter, each student enrolled in Team S will be required to participate either as an instructor or a "learner" in a weekly Minilesson Clinic. The procedural elements of that participation will be outlined below.

## One Booklet Required

During the first week of the quarter, when enrollments have settled down to some extent, a Minilesson Clinic assignment list will be distributed. Nine students will be assigned to a particular Minilesson Clinic group. Three of the students in each group will be designated as the A group, three as the B group, and three as the C group. (These arbitrary designations do not reflect the instructors' grading vision and will have no relationship to the final grade earned in the course.) Each student should, without delay, then go to the Student Book Store and purchase one copy of the Adult form Teaching Improvement Kit which bears the same letter as the one he or she was assigned. For example, a student in Minilesson Clinic No. 13 who is designated as a member of the B group should purchase a copy of Form B of the Adult Teaching Improvement Kit. Please note that there are some optional Teaching Improvement Kits available in the book store for use with children but that the required Teaching Improvement Kit, one of the three forms available (i.e., A, B, C), is an adult kit. With these materials in hand, the Minilesson Clinics will get underway the second week of the course in the assigned rooms.

One minilesson will be taught at each clinic, thus the individual assigned to teach on a given date must be present and prepared to teach. Failure to do so will result in a severe grade penalty. Yet, excused absences will occur. Therefore, all students should be ready to teach one week early in case of an unanticipated absence by the regularly assigned teacher.

## General Nature of the Participation

To provide a brief overview, an individual student's responsibilities will be described. First, a student will be obliged to instruct the six students in his Minilesson Clinic group who have been assigned other letters than his/her own. For example, a C member of the Minilesson group would be obliged to teach a lesson from the C Teaching Improvement Kit to the six A and B students in the group. Second, each student must act as a critic for two minilessons taught by the members of the minilesson group

who posses the same letters. To illustrate, a B member of Minilesson Clinic group No. 7 would not only teach one B minilesson him/herself, but would serve (along with another B member) as a critic while two other B lessons are being taught to the six A and C students. Finally, each student will be obliged to locate one or more groups of at least three adults to whom each of the three minilessons in his or her teaching kit can be taught, using a form of the posttest other than that employed in the Minilesson Clinic. Thus, in summary, each student will teach one minilesson to Team S classmates, critique two other minilessons being taught to Team S classmates, teach three minilessons to adults other than Team S classmates, and serve as a "learner" for six minilessons.

Procedural Specifics

Now, in more detail, here are the step-by-step procedures to be followed by each student. First, consult the minilesson assignment sheet to note which group, letter, and which week you have been assigned. Now read the first three chapters of the Teaching Improvement Kit (pp. 1-15). You may wish to examine the minilesson assignment information in Chapter 4. Do not examine the posttests (on green and blue sheets) which are included in the Teaching Improvement Kits. The minilesson clinic sessions have not been assigned to influence your grade in the course, but are specifically intended to help you improve your instructional skills. As a consequence, examination of the posttests prior to your teaching the lesson would reduce the likelihood that the minilesson would be beneficial to you.

On the day designated on the assignment sheet you will be obliged to teach one of the three lessons in your improvement kit (1, 2, or 3). On that day you will instruct (for no more than fifteen minutes) the six members of your group who have been assigned other letters, e.g., for the B students this would be the A and C students. The two members of your group who have been assigned the same letter will sit somewhat apart from the "student" group and will attempt to analyze the quality of your teaching plans, activities, etc. At the conclusion of the fifteen minute lesson the perforated copies of Posttest 1, (green paper) should be removed from your teaching improvement kits and distributed to your six "students." You should not have seen the posttest prior to this moment. These tests should be quickly completed and scored by the Minilesson Clinic group leader. On the basis of the average scores on (1) the interest rating and (2) the posttest, discussion of the teaching should be conducted in terms of the results produced. That is, good learner performance should result in a discussion focused on instructional tactics which seemed effective, poor results should lead to a discussion focused on instructional procedures which might be altered. To assist you in judging how effective a given teaching performance was, at the rear of each Teaching Improvement Kit data are available regarding how successful other teachers have been with each minilesson. Results of each student's performance will be turned in, but not for grading purposes, by the Minilesson Clinic leader.

On two occasions, therefore, each student will be serving not as a "learner" but as a critic. On those two occasions the student (having access to the Teaching Improvement Kit under consideration) will have had an opportunity to examine the minilesson assignment prior to the clinic. He or she will undoubtedly have some thoughts regarding an appropriate instructional procedure. This may be beneficial during the post-lesson analysis session. It will be useful to have two people, other than the Clinic leader, who are as conversant with the minilesson requirements as the teacher. During the analysis discussion it is anticipated that the Minilesson Clinic leader and the two critics will take primary responsibility for isolating elements of the lesson that were particularly effective or ineffective. Remember that the appraisal of instructional means should be made chiefly in terms of learner results, that is, the averaged interest ratings and posttest scores.

The final responsibility for each Team S student is to teach all three of the minilessons in his or her kit to another group of adults, that is, someone other than Team S students. Since each of the minilessons in the Teaching Improvement Kits has two equivalent posttests, the Team S students should not, either deliberately or inadvertently, examine Posttest 2 (blue sheets). Ideally after the original teaching of that minilesson in the Minilesson Clinic, each student should locate a group of at least three adults (friends, relatives, or people off the street) who would be willing to serve as students for the approximately 20 minutes involved for this assignment (15 minutes of teaching plus five or so minutes of testing). The minilesson should be taught and results summarized on the Minilesson Posttest 2 Summary Report Forms which have been provided (see attachment). The information called for is brief, yet should be completed in its entirety. These Minilesson Posttest 2 Report Forms should be turned in weekly during the 9 o'clock class meetings. In other words, an individual Team S student should have an opportunity to reteach the minilesson he originally taught to Team S students after a critique of the first lesson. Hopefully, this analysis will be useful in promoting improved learner performance. The other two lessons in the Kit will, for that student, be taught for the first time. Ideally the discussion of a classmate's teaching of that same lesson will prove useful in helping devise an effective lesson.

## Other Considerations

Results of the weekly Minilesson Clinics will be made available so that students who wish to compare their performance with that of other Team S members may do so. As indicated previously, there are other, optional Teaching Improvement Kits available in the book store for use with younger learners (elementary school age). Furthermore, minilessons in the adult kits can be used with mature secondary school students. Thus, it is possible to teach minilessons from another kit (other than the letter assigned) to high school age students. Either of these activities, that is, teaching minilessons to elementary or secondary school youngsters, should be considered optional for Team S students. Remember, the

whole purpose of the Minilesson Clinic operation is to improve the Team S member's skill in accomplishing prespecified objectives with teaching procedures which are also interesting to the learners. The amount of time that you can devote to promotion of this particular competency will undoubtedly yield great benefits to your future students. As a consequence, please approach the Minilesson Clinics and outside minilesson assignments as a real opportunity to increase your instructional skills.

MINILESSON POSTTEST 2 REPORT FORM

Your Name _____ Minilesson _____

Date Minilesson Taught _____ Today's Date _____

Setting for teaching and types of learners (One or two sentences)

_____

_____

_____

Number of students taught _____

Results: Average Interest Rating _____

Average Percent Correct on Posttest _____

Was this the second time you taught the minilesson? Yes ____ No ____

If this was the second time, supply the Average Interest Rating _____ and Average Percent Correct _____ for the first time you taught the minilesson.

Comments (Optional): _____

_____

_____

Below give the names and phone numbers (if available) of at least three of the students you taught.

_____

_____

_____